Mathematical Modelling, Vol. 2, pp. 179–190, 1981
Printed in the USA.

0270-0255/81/030179-12$02.00/0

ON SOME GUIDING PRINCIPLES IN MATHEMATICAL MODELLING WITH SPECIAL EMPHASIS ON DETERMINISM

DIEGO BRICIO HERNANDEZ

Departmento de Ingeniería
Universidad Autónoma Metropolitana (Iztapalapa)
Apdo Postal 55–534, México 13, D.F.

Communicated by X. J. R. Avula

Abstract—Various principles guiding the construction of mathematical models are isolated-selection of features to be considered in the model, character of time flow, character of space representation, determinism, causality, etc. It is shown how different choices of each of these items give rise to different classes of models—e.g., static vs dynamic, discrete time vs continuous time, lumped space vs distributed space, input–output systems vs stochastic systems, etc. Both causality and determinism are formalized in set theoretical terms. The validation problem of the models arrived at on deterministic-causal assumptions is dealt with along empiricist lines, the need for considering stochastic models emerging out of these considerations. A critical examination of determinism is given, out of which results the contention that determinism in a model merely expresses an attitude of the modeller towards the real world rather than an intrinsic property of a real system. Finally the conventional character of determinism is seen to be shared by the various other principles guiding the activity of the modeller, principles whose adoption is justified or invalidated only in terms of the comparison of the resulting model's predictions with actual observations.

1. METHODOLOGY OF MATHEMATICAL MODELLING

Whenever epistemological (science) or decisional (engineering) studies are undertaken, great care has to be exercised in securing good descriptions of the real system of interest. A *language* is required for that purpose, and mathematics has proven to be a very valuable aid in this respect. A *mathematical model* of a real system is a description of that system given within mathematics. Its elaboration requires a careful process of translation which consists of associating mathematical objects (i.e., sets and functions) with each of the relevant aspects of the system of interest.

There is no unique way of achieving that translation, the actual outcome being ultimately an expression of the tastes, feelings, previous experience, etc., of the translator (applied mathematician). However, the translator must endeavour to "preserve the structure" of the real system—the more so, the better—with the aim of obtaining faithful translations.

There is a group of principles guiding this activity. Each of them requires making decisions as to how time, space, and the various processes assumed to take place in space and time are to be represented in the description. These decisions refer to the following:

(a) TIME: Does it flow or remain static? If it does flow, is time flow continuous or discrete?
(b) SPACE: Is it distributed or lumped?

(c) PROCESSES: Which of them do take place? Is there any causal link between them? If so, is it deterministic or not?

The outcome of these decisions will dictate the choice of the various sets composing the structure

$$\langle T, S, A, E, \Omega \rangle, \tag{1.1}$$

where each of these sets is nonempty, and they have the following meanings:

(i) T represents *time*, and
   1. $T = \{0\}$ means time is regarded as static;
   2. $\{0\} \subsetneq T \subset \mathbf{Z}$ means time flows in a discrete fashion;
   3. T is an interval of real numbers when time flows in a continuous fashion.

(ii) S represents *space*, and
   1. "S is open in some Euclidean space of dimension $d \geq 1$" means space is distributed, of dimension d;
   2. "S is finite" means space is lumped.

In the distributed case, S inherits the Euclidean topological structure of $\mathbf{R}^d$, whereas it is regarded as a directed graph in the lumped case. In either case ∂S (the *boundary* of S) represents the interface system surroundings, the combined whole being represented by the union of S and its boundary, called the *closure* of S and denoted by $\bar{S}$.

(iii) A represents the values of the *attributes* of the system, whose spatial–temporal variation constitutes a *process*.

(iv) At each instant of time the attribute values are to be specified at each point of the closure of S, by means of a function $e : \bar{S} \to A$ to be called a *profile*. E consists of all *admissible profiles*, i.e., those possessing appropriate analytical properties.

(v) Ω consists of functions $\omega : T \to E$, $\omega(t)$ being "the profile at time t," for each $t \in T$, also to be denoted by ω_t. The extent of class Ω is a consequence of limitations imposed by our knowledge of the various processes taking place in the real system.

See [4] for a detailed illustration showing the construction of models like (1.1) under various assumptions, as well as a completion of such structure arrived at by adopting an empiricist viewpoint concerning the assignment of probabilities.

It must be noted that the classification of models implicit in the preceding description of the various elements of model (2.1) is far from being exhaustive. In fact, when modelling complex systems, it can happen that some parts of it are modelled with distributed space and some are not, some aspects are modelled deterministically and some other aspects are not, etc., so that in general "mixed" models have to considered.

The choice, however, is always open to the modeller, who is the only one to decide whether time is to be considered to flow or not, whether space is to be distributed or not, which processes are to be considered to actually take place, etc. The one and only limitation encountered by the modeller in this respect refers to the agreement of model predictions with actual observations on the real system: if considering a simple (say lumped space) model is not sufficient, the spatial distribution of the attributes should be considered instead; if assuming determinism does not lead to good predictions, try doing without that hypothesis, etc.

In general, trying the simpler alternative first seems to be a sensible modelling policy—one goes up the ladder of complexity if it is so required by the experiment anyway. It sometimes happens, though, that an already built model is to be simplified in order to extract conclusions from it more economically. This shows that "going down the ladder of complexity" is an interesting operation on models too. For a systematic approach to model simplification the reader is referred to the algebraic theory given by

Zeigler in [12]. The next section will concentrate on exploring the role played by determinism and causality in mathematical models.

2. CAUSALITY AND DETERMINISM IN MATHEMATICAL MODELS

Classically, the main problem associated with a model like that in (1.1) is that of "given present information plus whichever extra requirements are to be met, find future and past behaviour." Normally, present information is given in the form of an observation record over a certain time interval. Thus, a nonempty set $T_0 \subset T$ has to be specified, together with an observation record $\omega_0 \in E^{T_0}$. As to the extra requirements to be imposed, they normally refer to the action of the external world upon the real system under study and often can be embodied in a family $\{B_t, t \in T\}$ of sets of profiles, with the requirement that only those time evolutions ω should be considered for which $\omega_t \in B_t$ for each time instant $t \in T$. For the sake of illustration, suppose the real system is a metal rod one meter in length in which heat conduction is taking place. It is customary to take $T = \mathbf{R}$, $S = (0, 1)$ and $A = \mathbf{R}$, so that E consists of real functions e defined on $[0, 1]$ (temperature profiles). If both ends of the rod are to be kept at zero degrees, the subset B of E defined by

$$B = \{e \in E : e(0) = e(1) = 0\}$$

consists precisely of those profiles for which that happens, hence it is relevant. We can take $B_t = B$ for each $t \in T$ in this case, i.e., only those time evolutions ω for which $\omega_t(0) = \omega_t(1) = 0$ should be taken into consideration. In general, the classical problem of prediction can be posed in the context given by model (1.1) as follows:

Problem 2.1. Given $T_0 \subset T$, $\omega_0 \in E^{T_0}$ and $\{B_t, t \in T\}$, with $B_t \subset E$ for each $t \in T$, find $\omega \in \Omega$ such that

(i) $\omega|_{T_0} = \omega_0$

(ii) $\omega_t \in B_t$ for each $t \in T$.

Numerous problems in classical science and engineering can be cast in this fashion, normally in terms of the differential or difference equations specifying Ω, case in which (i) is known as the *initial condition* (and normally $T_0 = \{t_0\}$), whereas (ii) constitutes the *boundary condition* in most cases of interest.

Let

$$\Omega_0 = \{\omega \in \Omega : \text{(i), (ii) hold}\},$$

i.e., Ω_0 is the solution set of Problem 2.1. If $\Omega_0 = \emptyset$, requirements (i) and (ii) are obviously excessive, and a logical analysis aimed at establishing the reasons for that incompatibility is in order. If Ω_0 consists of only one element (i.e., an existence and uniqueness theorem has been proved for Problem 2.1) then a "sensitivity analysis" is carried out as follows: let $\omega', \omega'' \in \Omega$ be the (unique) solutions of Problem 2.1 corresponding to data $\omega_0', \{B_t'\}$ and $\omega_0'', \{B_t''\}$, respectively. If the two sets of data "differ little," will the corresponding solutions "differ little" too? The notion of "differing little" can be formalized in topological terms—i.e., distance—which transforms the question just posed into one referring to the continuity properties of the map

$$(\omega_0, \{B_t\}) \mapsto \omega,$$

i.e., continuity of the solution with respect to data. Then, if the last question can be

answered in the affirmative, Problem 2.1 is said to be *well posed*, a term originally coined by J. Hadamard in connection with partial differential equations [2]. This property of being well posed is highly significant, since it shows that the model is well behaved in the following sense:

In practice, both initial and "boundary" conditions are known only approximately; continuity of solutions with respect to initial data means that the errors in the determination of the system's behaviour can be made as small as desired provided the errors in the data are kept sufficiently small. On the other hand, existence and uniqueness of solutions is a necessary prerequisite for even posing the question of continuity of solutions with respect to data. Moreover, it reflects a property frequently observed in real systems: there is only one behaviour compatible with initial data once the side conditions have been specified. This observation is the essence of *determinism*.

Lastly, if Ω_0 consists of more than one element (i.e., there is existence of solutions but uniqueness has been proved not to hold) then conditions (i) and (ii) in Problem 2.1 do not completely specify future behaviour, and this may indicate that the model is incomplete. This question will be given ample coverage in the next two sections, where model (1.1) will be appropriately supplemented in order to deal with nondeterministic situations, and later to review critically the notion of determinism.

For the time being it will prove very useful to introduce the notion of *causality*, which will reinforce that of determinism in a very natural way. For suppose an existence and uniqueness theorem can be proved for Problem 2.1. It may happen that by fixing certain features of the model (say, by specifying the initial data), for every choice of time evolution of a certain part of the profiles (like boundary conditions, for example) there is one and only one time evolution for the remaining part of the profiles. This idea can be illustrated in more concrete terms by referring to a tank into which a liquid is being discharged at a certain rate, and out of which flow its contents through an orifice at the bottom: given the initial quantity of water in the tank, for every choice of feed rate there is only one possible time behaviour of the amount of liquid in the tank.

In general, it is sometimes possible to identify E with the Cartesian product of two sets U and Y,

$$E = U \times Y,$$

so that E-valued functions are ordered pairs composed of a U-valued function and a Y-valued function, i.e.,

$$E^T = U^T \times Y^T.$$

It is convenient to define

$$\Gamma = \{\gamma \in U^T : (\gamma, \eta) \in \Omega \text{ for some } \eta \in Y^T\}$$
$$H = \{\eta \in Y^T : (\gamma, \eta) \in \Omega \text{ for some } \gamma \in U^T\}.$$

Then each $\omega \in \Omega$ can be written as

$$\omega = (\gamma, \eta)$$

for suitable $\gamma \in \Gamma$, $\eta \in H$. In other words,

$$\Omega \subset \Gamma \times H.$$

Moreover, it is sometimes possible to find a nonempty set X—somehow associated with model (1.1)—with the following property:

Causality 2.2. Given $x \in X$, for each $\gamma \in \Gamma$ there is one and only one $\eta \in H$ such that $(\gamma, \eta) \in \Omega$.

Then, there is said to be a *causal link* between those parts of the profiles contained in U and those contained in Y; equivalently, Γ is said to be the *cause* of H, and H is said to be the *effect* of Γ. In the usual systems theoretic terminology, the elements of Γ are called *inputs*, those of H are called *outputs*, and those of X (initial) *states* [12, p. 205] or *experiments* [5, p. 11]. Causality is equivalent to the existence of a function $f_x : \Gamma \to H$ for each $x \in X$, with $f_x(\gamma)$ = output obtained when input γ is fed into the system if its initial state was x.

Of particular interest is the causal relation in which the effect does not anticipate the cause. In fact, this lack of anticipation is what ultimately enables us to define an order in time (see [9, §21]). This feature of causality can be reflected by the functions mapping inputs into outputs if each of them is nonanticipative in the sense of the following definition, in which

$$T_t^s = \{r \in T : s \leq r \leq t\} \tag{2.3}$$

is the time interval between s and t. One writes T^t and T_t instead of T_∞^t and $T_t^{-\infty}$, respectively.

Definition 2.4. Let $s \in T$, $\varphi : \Gamma \to H$ be given. φ is said to be *s-nonanticipative* if, for each $t \in T^s$,

$$\eta'(t) = \eta''(t) \text{ if } \gamma'|_{T_t^s} = \gamma''|_{T_t^s} \tag{2.4}$$

whichever inputs $\gamma', \gamma'' \in \Gamma$ are chosen, with $\eta' = \varphi(\gamma')$, $\eta'' = \varphi(\gamma'')$. φ is *nonanticipative* if it is $-\infty$-nonanticipative.

In other words, under a nonanticipative f_x the effect at a given instant depends only on the cause prior to that instant, with no dependence whatsoever on its future.

Therefore, under causality, model (1.1) specializes into the structure

$$\langle T, U, \Gamma, Y, H, \{f_x, x \in \mathbf{X}\}\rangle, \tag{2.5}$$

which is called a *dynamical* (*input–output*) *system* [5, p. 10], if each f_x is nonanticipative. Family $\{f_x, x \in \mathbf{X}\}$ is termed the *input–output relation* of the system. Observe that a causal model (input–output system) is deterministic in the sense that once the initial state is fixed there is one and only one output for each choice of input, hence only one trajectory of the system.

3. MODEL VALIDATION

The purpose of modelling reaches well beyond constructing a structure like (1.1) or (2.5): it aims at choosing one such model that faithfully reproduces observed behaviour, i.e., a "valid" model. In short, our point of view will be *empiricist*, in the sense that a model can be validated only through successful comparison of its predictions with experiment. The following paragraphs are aimed at rendering precise this notion, i.e., that of a model like (2.5) being "successfully compared with experiment."

Classically (e.g., see [12, p. 195]) a model like (2.5) will be regarded as *valid* if the

following situation holds for every choice of *set of observation instants* T_1, *test input* γ, and corresponding *observation record* η_1:

Let $T_1 \subset T$ be given, its elements representing those instants at which output values are measured. Pick $\gamma \in \Gamma$ and let $\eta_1 \in Y^{T_1}$ be the output values recorded during the experiment in which the real system is subjected to input γ. Then, there is an initial state $x \in X$ such that

$$\eta \mid_{T_1} = \eta_1 \tag{3.1'}$$

with

$$\eta = f_x(\gamma). \tag{3.1''}$$

Note, however, that this validation criterion is rather strict, and in practice it is often replaced by one of *approximate validation* in which equality in (3.1) is replaced by "being approximately equal." This concept can be formalized in terms of a *tolerance* given in the form of a positive number ρ, η_1 and $\eta \mid_{T_1}$ being *approximately equal* if they *differ by less than* ρ, this last notion being in turn formalizable in terms of a distance function d_1 defined on Y^{T_1}. This gives rise to the condition

$$d_1(\eta_1, \eta \mid_{T_1}) < \rho \tag{3.2}$$

as the counterpart of (3.1) for approximate validation. Needless to say, both ρ and d_1 are supplied by the modeller.

Moreover, it happens in practice that successive repetitions of the same experiment "measuring the system's response to input γ over the time interval T_1" have different outcomes, say $\eta_1^{(1)} \ldots \eta_1^{(N)}$, if N denotes the number of repetitions of the same experiment. It should not be surprising then that often not all of the "approximate equality conditions"

$$d_1(\eta_1^{(1)}, \eta \mid_{T_1}) < \rho \ldots d_1(\eta_1^{(N)}, \eta \mid_{T_1}) < \rho$$

will hold. However, if k of them happen to be true and "k/N is close to unity," say

$$1 - \epsilon \leq k/N \leq 1 \tag{3.3'}$$

with ϵ given beforehand (and "small," as a rule), then the discrepancy will be attributed to "experimental error" and henceforth ignored. Quotient k/N is referred to as the *relative frequency* of occurrence of the validation condition (3.2) during the N repetitions of the experiments, and will be denoted by f_N, i.e.,

$$f_N = \frac{k}{N}.$$

But, what happens if (3.3') does not hold, as is sometimes the case? Again, f_N is "small," say

$$0 \leq f_N < \epsilon, \tag{3.3''}$$

then there should be no objection to rejecting the approximate equality (3.2) of model prediction and observation record over T_1, hence the model should be regarded as

invalid. Again, the choice of ϵ in an actual situation falls under the sole responsibility of the modeller.

But, what if f_N is neither close to zero nor close to unity, say $\epsilon \leq f_N < 1 - \epsilon$? Again, considerations connected with experiment give us a clue as to what to do. For, it is an often confirmed fact that as the number N of repetitions of the same experiment grows, the relative frequency fluctuates about a number $\alpha \in [0, 1]$, the amplitudes of the fluctuations decreasing as more and more repetitions of the same experiment are considered. This observation embodies a generalization from empirical evidence, and forms the basis for the assignation of probabilities from an empiricist position (see [8]). Retaking our main discussion, a sensible thing to do is to associate this value α with the validation condition (3.2), or rather with the set of output trajectories that satisfy it, namely,

$$C = \{\eta' \in H : d_1(\eta'|_{T_1}, \eta|_{T_1}) < \rho\}. \tag{3.4}$$

On a strict basis, a model is *valid* if and only if for each T_1 and for each $\gamma \in \Gamma$ there is an $x \in X$ such that α is "close to 1"; it is *invalid* otherwise, i.e., if there is a choice of T_1 and γ such that for no choice of initial state x, α happens to be "close to 1." Here the phrase "being close to 1" has to be given a meaning, a possibility being "lying between $1 - \epsilon$ and 1," as in (3.3').

Note that, strictly speaking, the validation of a model according to this scheme requires conducting "a large number" N of experiments for each choice of T_1 and γ, plus being able to find the right initial state for each such choice. This is rather difficult in practice, and actually impossible to achieve if either T or Γ is an infinite set, since only a finite number of experiments can at most be expected to be performed in any situation. Thus, it is easy to convince ourselves that the decision of labelling a model as valid or invalid will have to be made on the basis of restricted information, and in any case will involve an appeal to *induction.*

Moreover, the actual technique presented here is rather simple minded, it being aimed not at furnishing a practical validation procedure, but rather at providing a conceptual framework for model validation in terms as simple as possible. In practice, model validation requires the concourse of statistical techniques of various kinds (notably hypothesis testing and parameter estimation), our simple minded consideration of "limiting values of relative frequencies" being an elementary instance of a statistical notion.

Note, however, a very fundamental point: it suffices to show the slightest regard for experimental results to make the modeller resort to nondeterministic considerations in order to validate a deterministic model. In fact, no modeller would seriously attempt to validate a deterministic model like (2.5) by following the dictates of the validation criterion centered around (3.1) or even (3.2), certainly not if he is willing to face the fact that repetitions of the same experiment do not always give the same results. By so doing, the modeller should learn the lesson that the class of deterministic models is much too restricted. A meaningful, self-consistent validation theory requires a larger class of models, stochastic models providing one such choice.

4. STOCHASTIC MODELS

Suppose no valid deterministic model has been found for a certain real system. Whether such a state of affairs should be regarded as being peculiar to the real system being considered or simply as a measure of our ignorance of such system will not be dwelled upon for the time being—see Sec. 6, however. In such a case, a model like (1.1)

will have to be considered in which Problem 2.1 cannot be solved uniquely, namely one in which determinism is not reflected.

Now, observe that determinism can be subsumed into causality as given in statement 2.2 provided one can take

$$\mathbf{X} = E^{T_0}$$

and, assuming

$$E = U \times Y$$

hence

$$\Omega \subset \Gamma \times H$$

as before, set

$$B_t = \{\gamma(t)\} \times Y$$

in Problem 2.1 for fixed $\gamma \in \Gamma$. In this setting, the fact that model (1.1) does not reflect determinism specializes into X not qualifying as "set of initial states." In other words, there are "initial data" $x \in X$ and inputs $\gamma \in \Gamma$ for which there are at least two *distinct* elements ω', ω'' of Ω such that

$$\omega' |_{T_0} = x = \omega'' |_{T_0}$$

and both conditions

$$\omega'_t = (\gamma(t), y') \text{ for some } y' \in Y$$

$$\omega''_t = (\gamma(t), y'') \text{ for some } y'' \in Y$$

hold for each $t \in T$.

Again rephrasing the same statement, for such a set of initial data x the assignment $\gamma \mapsto \eta$ is *not* a function, contrary to the statement of 2.2. What could replace the functions present in the input–output relation of model (2.5) arrived at under the assumption of deterministic causality?

At this point the approach adopted in the previous section will prove to be very fruitful. For, observe that not only conditions such as (3.2) can be checked through experiment. In general, a set of "output records" $B \subset Y^{T_1}$ can be given, and whether

$$\eta |_{T_1} \in B \tag{4.1}$$

can be checked in exactly the same fashion in every repetition of the experiment. Note that (4.1) specializes to (3.4) if

$$B = \{\eta' \in Y^{T_1} : d_1(\eta', \eta |_{T_1}) < \rho\}.$$

In turn, condition (4.1) defines a set of output trajectories

$$C = \{\eta \in H : \eta |_{T_1} \in B\}, \tag{4.2}$$

a set that can be assigned a probability by repeating the experiment in question a "large number" of times and identifying the limiting value $\alpha \in [0, 1]$ of the relative frequency with which condition (4.1) has occurred. Thus, every set of the form (4.2)—usually referred to as a *cylinder set* in H—can be assigned a probability through experiment.

This observation constitutes a clue for the construction of models from empirical data. In fact, if $\mathscr{C}$ denotes the class of all cylinder sets of H, the experiment provides (at least conceptually) an assignation of probabilities $C \mapsto \alpha$ defined on the whole of $\mathscr{C}$. Recalling that this assignation is made for fixed $\gamma \in \Gamma$, then a stochastic model results by associating with every $\gamma \in \Gamma$ the corresponding assignment of probabilities to the cylinder sets in H. This correspondence constitutes the required generalization of the input–output relation of (2.5) to a nondeterministic situation.

Under certain conditions, for each input $\gamma \in \Gamma$ the corresponding probabilization of the cylinder sets in H can be extended to a larger class $\mathscr{H}$ that contains $\mathscr{C}$, in such a way that the resulting probabilization of the sets in $\mathscr{H}$ can be called a *probability* in the technical sense of the term (see [7, 1]). A very important result pointing definitely in the direction of asserting the feasibility of this empiricist approach is a theorem due to the Soviet mathematician A. N. Kolmogoroff, the much celebrated *consistency theorem* [7]. See Sec. 3 of [4] in this connection.

Thus, a stochastic model is arrived at by assigning a probability distribution on the output space H to every input γ. This correspondence reduces to an H-valued function when each of the probability distributions associated with the inputs is concentrated at a single point γ of H —the output corresponding to the input being considered. Thus, the class of stochastic models contains the class of input–output systems as a subclass. The reader is referred to Sec. 3 of [3] for a technical presentation of these ideas, one that incorporates an extension of the notion of nonanticipativity in causality that is applicable in this more general context.

5. CONVENTIONALISM IN MODELLING

This section responds to the need of placing in a proper perspective the various modelling principles referred to in Sec. 1. To begin with, the principle of determinism will be examined in the light of the following examples:

Example 5.1. Consider the free fall of a coin from a certain height. Three aspects of this phenomenon will be isolated here: the initial position of the coin, its free fall time, and the side showing upwards once the coin is in rest. Note that the first of these three aspects can be externally controlled, as opposed to the remaining two which can only be observed. On the other hand, experiment gives results such as

Repetition	1	2	3	4	5
Time (sec)	1.80	1.85	1.75	1.80	1.80
Side	1	0	0	1	1

corresponding to the experiment in which a coin is released from a point 16 m above the ground, with side = 1(0) if heads (tails) comes up.

Experiments similar to the one just described give rise to the suspicion that for a given initial height the falling time can be predicted with great accuracy. In fact, the average value of the five values of falling time recorded in Table 1 is 1.80 sec, with a maximum deviation of 2.88% with respect to this value. This is a small figure, so small that we feel tempted to neglect it and assume instead that "falling time is a function of height," the observed variability being explained away in terms of "experimental error."

Experimenting with various height values yields the law that "the falling time is

proportional to the square root of the initial height." This piece of information suggests that a reasonable model could be cast in the shape of an input–output system with

$$T = \{1\},\ U = Y = [0, \infty),\ \Gamma = U,\ H = Y,\ \mathbf{X} = \{1\}$$

and input–output function $f : U \to Y$ given by

$$f(u) = k\sqrt{u}$$

with $k > 0$ and $f(u)$ = falling time when coin is released from height u. This model is easily validated in practice, provided the proportionality constant k is chosen appropriately ($k = 0.452$ sec/m^2 for experiments conducted at sea level).

On the other hand, consider the other aspect of this phenomenon ("heads or tails") or rather its dependence with height. If it is to be modelled by an input–output system, it must have

$$Z = \{0, 1\}$$

as output space and an input–output function $g : U \to Z$, all the remaining ingredients in

$$\langle T, U, \Gamma, Z, Z, g \rangle$$

retaining their meaning. Noting that

$$U = g^{-1}(0) + g^{-1}(1),$$

such a model would predict that the coin definitely falls heads when dropped from certain well-specified heights, the outcome being invariably tails in any other case. As already noted, these predictions are far from being confirmed by experiment hence a model of this type could not possibly be validated under reasonable criteria. In short, we are led to conclude that this side of the phenomenon being analyzed is nondeterministic. A stochastic model for it consists essentially of a probability on Z, namely,

$$\emptyset \mapsto 0,\ \{1\} \mapsto p,\ \{0\} \mapsto 1 - p,\ Z \mapsto 1$$

with p = probability of heads.

The conclusion to be drawn from these results seems to be that there is a causal deterministic link from height to falling time, whereas no such situation arises if "falling time" is replaced by "side," at least not in general. Note, however, that p can be found to be "very close to 1" for a particular coin (a highly biased one), a case in which a deterministic model with $g(u) = 1$ for all $u \in U$ could well be validated. On the other hand, recall that the existence of an input–output function like f or g was validated on the grounds that "the relative deviation of the results with respect to their average value was very small," yet that variability in the results simply could not be ignored under certain conditions.

For example, suppose the experiments are carried out for small heights only, say a few centimeters, time being measured using a not very precise clock, so that precision to within one second simply cannot be guaranteed. Possible results for the experiment in which the coin is released from a height of 50 cm are as follows:

Repetition	1	2	3	4	5
Time (sec)	1	0	0	1	1

Thus, under these conditions, no deterministic model could be reasonably validated for this aspect of the phenomenon being considered.

Example 5.2. Suppose we are interested in studying radioactive disintegrations. If the investigation is carried out on a "long term" basis (say hours) the experimental facilities could very well consist of a watch and a scale, weight losses becoming appreciable for this duration of the experiment. This is not so when this phenomenon is studied on a short term basis (say a few seconds), this situation requiring a good chronometer and a Geiger counter. Experiment shows that this phenomenon appears distinctly deterministic when viewed on a long term basis, determinism vanishing when studied in the second experimental setup. See Sec. 2 of [3] for a more detailed presentation of these considerations.

These two examples indicate that the same phenomenon may appear to be deterministic when examined from a certain perspective, and yet lose this characteristic when viewed differently. Moreover, both observer and observation scale seem to play a prominent role in this assessment.

This observation leads us to conclude that determinism is *not* an attribute that can safely be said to be possessed by certain parts of the real world and not by certain other portions of it. It is rather a property of the (mathematical) model used to describe the real system, a property resulting from basic assumptions made by the modeller on the phenomenon in question, and whose justification or invalidations comes ultimately from the confrontation between model predictions and experimental results. This point of view concerning determinism is akin to the *conventionalistic* approach concerning the axioms of geometry developed by Poincaré (see [10, p. 50], for example) and can be said to supplement it.

This conventional character is by no means peculiar to this modelling principle only. Take the choice of time set in (1.1), for example. Surely a basic requirement on any candidate for T is that it should be an ordered set, since otherwise it would not be possible to speak of "past," "present" and "future" in terms of it. The requirement imposed by the continuity of time flow seems to be that "there should be no consecutive events" or, in other words, "between any two events there conceivably occurs a third event." An interval of rational numbers could well qualify as time set under these conditions, no doubt, yet intervals of real numbers are chosen instead simply on account of the more convenient analytical properties of $\mathbf{R}$, i.e., the possibility of taking limits. On the other hand, given the precision offered by most time measuring devices, no empirical arguments of any kind could be used to justify the need for including irrationals in T. In fact, for a given accuracy, only finitely many rational time values will ever have to be considered, so time might just as well be regarded to flow in a discrete fashion, thus erasing the dividing line between both types of time flow. It seems to be that, in the last resort, the type of time flow incorporated in a model results from considerations of convenience only. Moreover, why should time be assumed to flow at all? Why not consider a static model instead? Again, a certain phenomenon can be safely regarded to be static when examined in a "short-term" basis, yet considering the same phenomenon on a "long-term" basis may require assuming time to flow if a good model is to be arrived at. Needless to say, both "short" and "long" have to be made precise at each modelling instance, this requiring an appeal to experience. Again we have to face another modelling principle bearing a conventional character: it is only the confrontation of a model's predictions with experiment that validates or invalidates a certain choice for T in (1.1), and a valid choice may become invalid if the experimental conditions are changed.

Similar arguments can be used to show that the choice of all the other elements in (1.1) obeys equally conventional modelling principles. These points are illustrated in a particular context in [4], where various choices of space and time representation are illustrated.

It appears convenient to conclude at this stage by pointing out two types of consideration usually made when choosing the various elements making up a model like (1.1):

(a) *Purpose of the model.* Once the attributes of interest have been identified and sets A and S have been chosen, the choice of E often results from reasons of analytical convenience. However, its factorization into factors U and Y results from identifying those parts of the profiles that can be acted upon externally (U) and those that can only be observed (Y). Different experimental setups would obviously manifest in different choices for these sets.

(b) *State of knowledge.* The amount of knowledge the modeller has on the real system imposes restrictions on the time evolution of the profiles. It makes the modeller require certain relations to hold (e.g., conservation laws, constitutive relations, etc.), thus dictating the choice of Ω. Modellers differing in this respect might choose differently.

Acknowledgements—The publication of this material owes a great deal to the sustained interest of Dr. Nicola Bellomo (Politecnico di Torino). Many of the ideas expressed here have been discussed with the Automation group of Instituto de Ingenieria (UNAM). My deepest appreciation to all of them. Italian version of this paper appeared in *Quale Scenza* (Torino, 1980).

REFERENCES

1. M. Fréchet, Las Matemáticas y lo Concreto, in *Col. Problemas Cientificos y Filosóficos*, 10th Ed., UNAM, Cd. de México (1958).
2. J. Hadamard, *La Theorie des Equations aux Derivées Partielles*, Editions Scientifiques, Peking (1964).
3. D. B. Hernandez, Towards a Conventionalist Empiricist View of Mathematical Modelling, in *Proceedings of the First International Conference on Mathematical Modelling*, edited by X. J. R. Avula, Vol. I, pp. 381–391, St. Louis, MO (1977).
4. D. B. Hernandez, Science and Mathematics: Prediction Models (in Spanish). *Rev. Ciencia* **31**, 103–121 (1980).
5. R. Kalman, P. L. Falb, and M. A. Arbib, *Topics in Mathematical Systems Theory*, McGraw-Hill, New York (1969).
6. J. L. Kelley, *General Topology*, Van Nostrand Reinhold, New York (1955).
7. J. Neveu, *Bases Mathématiques du Calcul des Probabilités*, Masson, Paris (1970).
8. H. Reichenbach, *The Rise of Scientific Philosophy*, University of California, Berkeley (1951).
9. H. Reichenbach, *The Philosophy of Space and Time*, Dover, New York (1958).
10. H. Poincaré, *Science and Hypothesis*, Dover, New York (1952).
11. P. Suppes, *Introduction to Logic*, Van Nostrand, Princeton, NJ (1957).
12. B. P. Zeigler, *Theory of Modelling and Simulation*, Wiley, New York (1975).